Amplification of a complete simian immunodeficiency virus genome from fecal RNA of a wild chimpanzee

JOURNAL OF VIROLOGY, Feb. 2003, p. 2233–2242 Vol. 77, No. 3
0022-538X/03/$08.00+0 DOI: 10.1128/JVI.77.3.2233–2242.2003

Amplification of a Complete Simian Immunodeficiency Virus Genome from Fecal RNA of a Wild Chimpanzee

Mario L. Santiago,[1] Frederic Bibollet-Ruche,[1] Elizabeth Bailes,[2] Shadrack Kamenya,[3] Martin N. Muller,[4] Magdalena Lukasik,[3] Anne E. Pusey,[5] D. Anthony Collins,[3] Richard W. Wrangham,[4] Jane Goodall,[6] George M. Shaw,[1,7] Paul M. Sharp,[2] and Beatrice H. Hahn[1]*

Departments of Medicine and Microbiology, University of Alabama at Birmingham,[1] and Howard Hughes Medical Institute,[7] Birmingham, Alabama 35294; Institute of Genetics, University of Nottingham, Queens Medical Centre, Nottingham NH7 2UH, United Kingdom[2]; Gombe Stream Research Centre, The Jane Goodall Institute Tanzania, Kigoma, Tanzania[3]; Department of Anthropology, Harvard University, Cambridge, Massachusetts 02138[4]; Department of Ecology, Evolution and Behavior, University of Minnesota, St. Paul, Minnesota 55108[5]; and The Jane Goodall Institute, Silver Spring, Maryland 20911[6]

Received 11 September 2002/Accepted 30 October 2002

Current knowledge of the genetic diversity of simian immunodeficiency virus (SIVcpz) infection of wild chimpanzees (*Pan troglodytes*) is incomplete since few isolates, mostly from captive apes from Cameroon and Gabon, have been characterized; yet this information is critical for understanding the origins of human immunodeficiency virus type 1 (HIV-1) and the circumstances leading to the HIV-1 pandemic. Here, we report the first full-length SIVcpz sequence (TAN1) from a wild chimpanzee (*Pan troglodytes schweinfurthii*) from Gombe National Park (Tanzania), which was obtained noninvasively by amplification of virion RNA from fecal samples collected under field conditions. Using reverse transcription-PCR and a combination of generic and strain-specific primers, we amplified 13 subgenomic fragments which together spanned the entire TAN1 genome (9,326 bp). Distance and phylogenetic tree analyses identified TAN1 unambiguously as a member of the HIV-1/SIVcpz group of viruses but also revealed an extraordinary degree of divergence from all previously characterized SIVcpz and HIV-1 strains. In Gag, Pol, and Env proteins, TAN1 differed from west-central African SIVcpz and HIV-1 strains on average by 36, 30, and 51% of amino acid sequences, respectively, approaching distance values typically found for SIVs from different primate species. The closest relative was SIVcpzANT, also from a *P. t. schweinfurthii* ape, which differed by 30, 25, and 44%, respectively, in these same protein sequences but clustered with TAN1 in all major coding regions in a statistically highly significant manner. These data indicate that east African chimpanzees, like those from west-central Africa, are naturally infected by SIVcpz but that their viruses comprise a second, divergent SIVcpz lineage which appears to have evolved in relative isolation for an extended period of time. Our data also demonstrate that noninvasive molecular epidemiological studies of SIVcpz in wild chimpanzees are feasible and that such an approach may prove essential for unraveling the evolutionary history of SIVcpz/HIV-1 as well as that of other pathogens naturally infecting wild primate populations.

West-central African chimpanzees (*Pan troglodytes troglodytes*) are naturally infected with strains of simian immunodeficiency virus (SIVcpz) that represent the closest relatives of human immunodeficiency virus type 1 (HIV-1) groups M, N, and O and have thus been implicated as the primate source of the human infections (12). East African chimpanzees (*Pan troglodytes schweinfurthii*) are also infected with SIVcpz, but genetic information for these viruses is limited since only two such strains (ANT and TAN1) have so far been identified (32, 45). The first of these, SIVcpzANT, was isolated over a decade ago from a wild-caught chimpanzee orphan (Noah) which was confiscated upon illegal exportation from Kinshasa (Democratic Republic of Congo) (26). Although its geographic origin is unknown, this chimpanzee was subsequently typed as an eastern chimpanzee (*P. t. schweinfurthii*) by mitochondrial DNA analysis (12). SIVcpzANT was completely sequenced and found to be most closely related to members of the SIVcpz/HIV-1 group of viruses; however, phylogenetic analysis revealed that it formed a highly divergent outgroup to all previously characterized SIVcpz and HIV-1 strains (45). The second *P. t. schweinfurthii* virus, SIVcpzTAN1, was discovered only recently in a wild chimpanzee (Ch-06) from Gombe National Park in Tanzania (32). Using noninvasive detection methods, we identified SIVcpz-specific antibodies in urine samples and confirmed SIVcpz infection by amplifying viral sequences from fecal RNA. Again, sequence analysis indicated infection with a highly divergent SIVcpz strain (32); however, the availability of only partial *pol-vif* sequences precluded a conclusive phylogenetic analysis. In this paper, we describe a novel reverse transcription-PCR (RT-PCR) strategy allowing the amplification of the complete SIVcpzTAN1 sequence from fecal samples of chimpanzee Ch-06. Analysis of this sequence revealed new insight into the evolutionary history of the HIV-1/SIVcpz group of viruses and provided conclusive evidence for the existence of two divergent

* Corresponding author. Mailing address: Department of Medicine, University of Alabama at Birmingham, 720 20th St. South, Kaul 816, Birmingham, AL 35294. Phone: (205) 934-0412. Fax: (205) 934-1580. E-mail: bhahn@uab.edu.

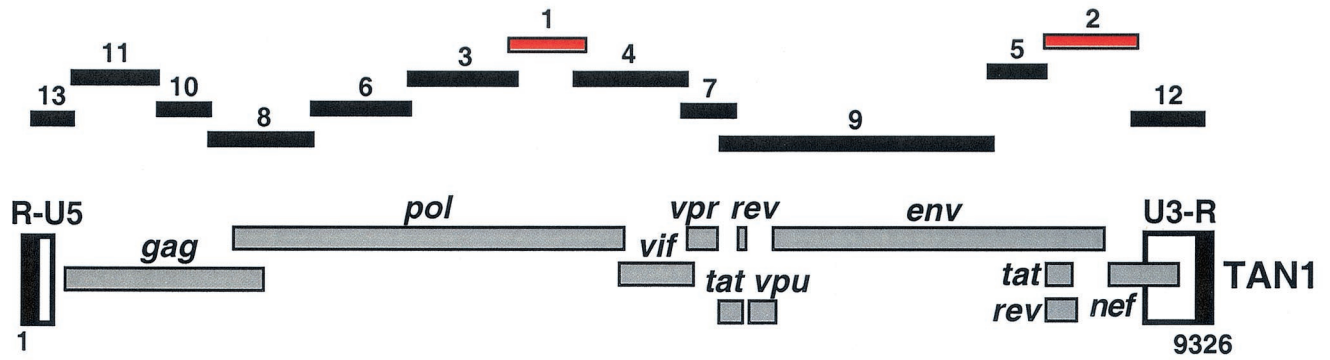

FIG. 1. Amplification of a complete SIVcpzTAN1 sequence from chimpanzee fecal RNA. RT-PCR was used to amplify partially overlapping subgenomic fragments from virion RNA extracted from two independent fecal samples collected on 9 November 2000 and 11 November 2000 from chimpanzee Ch-06 under direct observation. The relative positions of the fragments with respect to the SIVcpzTAN1 genome, along with their order of amplification, are shown. Fragment lengths are drawn to scale (the entire SIVcpzTAN1 sequence is 9,326 bp in length). The first two fragments, which were derived with consensus primers from highly conserved regions of the HIV-SIV genome, are highlighted in red. Sequences of fragments 1, 3, and 4 have been reported previously (32).

SIVcpz lineages infecting chimpanzees in west-central and east Africa, respectively.

Chimpanzee Ch-06 was a healthy, 24-year-old, sexually active, midranking male member of the Kasekela community, which comprises approximately 55 individuals and is the largest in Gombe National Park. All community members live freely but are habituated to the presence of human observers (14). They are individually known and named and are routinely studied at close distances. Fecal samples were collected from chimpanzee Ch-06 on three independent occasions (9 November 2000, 11 November 2000, and 18 November 2000) under direct observation by field researchers from the Gombe Stream Research Centre. (Fieldwork in Gombe National Park was authorized by the Tanzania Commission for Science and Technology, the Tanzania Wildlife Research Institute, and the Tanzania National Parks.) Approximately 20 g of fecal material was placed into a 50-ml tube containing 20 ml of RNA*later* (Ambion, Austin, Tex.). RNA*later* preserves nucleic acids, allowing storage and shipment at room temperature without viral RNA degradation. Tubes were sealed with Parafilm, initially stored at −20°C, and then shipped at ambient temperature to the United States by air carrier. Upon arrival, fecal suspensions were aliquoted and stored at −70°C, and RNA was extracted with the RNA*queous*-Midi kit (Ambion) according to the manufacturer's recommendations. Briefly, fecal aliquots (1 ml) were resuspended in lysis buffer (6 ml), vortexed, and centrifuged (16,000 × *g*; 3 min) to remove solid debris. Clarified supernatants were mixed with an equal volume of 64% ethanol and passed through a glass fiber filter unit to bind nucleic acids. Following extensive washes, bound nucleic acids were eluted (1 ml), precipitated with $LiCl_2$ to enrich for RNA and to remove potential PCR inhibitors, pelleted, and resuspended in 50 μl of distilled water. RNA extracts were stored at −70°C.

Previous studies of HIV-1- and SIVcpz-infected chimpanzees as well as HIV-1-infected humans had shown that viral RNA, but not proviral DNA, was detectable in fecal samples, suggesting that viral nucleic acids were present in virions and not in infected cells (32, 46). We thus used an RT-PCR approach with consensus and strain-specific primers to amplify partially overlapping subgenomic (400- to 2,000-bp) viral fragments in a stepwise fashion to derive a complete SIVcpz sequence (Fig. 1). For cDNA synthesis, 10 μl of extracted fecal RNA was added to an RT-PCR master mix consisting of 1× buffer II (Perkin-Elmer, La Jolla, Calif.), 5 mM $MgCl_2$, 1 mM deoxynucleoside triphosphate, 5 mM dithiothreitol, 20 pmol of reverse primer (Table 1), 20 U of RNase inhibitor (Promega, Madison, Wis.), and 100 U of Superscript RT II (Gibco-BRL, Rockville, Md.) and incubated for 1 h at 42°C. Ten microliters of this cDNA preparation was then added to a PCR mix consisting of Expand Buffer 2 (Roche Molecular Biochemicals, Indianapolis, Ind.), 0.35 mM deoxynucleoside triphosphate, 2.25 mM $MgCl_2$, 0.1 μg of bovine serum albumin/ml, 2.5 U of Expand high-fidelity *Taq* polymerase (Roche Molecular Biochemicals), and 10 pmol of forward and reverse primers (Table 1) and subjected to nested (or seminested) PCR amplifications. Amplification conditions included 45 cycles of denaturation (94°C, 0.5 min), annealing (50°C, 0.5 min), and elongation (68°C, 1.5 min). Two microliters of the first-round reaction mixture was used for second-round amplification under the same thermocycling conditions.

Targeting regions of high sequence conservation, we initially amplified a 650-bp fragment in the integrase region of *pol* and a 782-bp fragment spanning the *gp41-nef* junction (fragments 1 and 2 in Fig. 1). SIVcpzTAN1 strain-specific primers were then combined with upstream and downstream consensus primers to obtain sequences 3′ and 5′ of these anchor fragments (Table 1). Fragments of 1,000 bp or less were readily amplified. However, attempts to amplify longer fragments were generally unsuccessful, except for a 2,197-bp envelope fragment which was obtained with SIVcpzTAN1-specific primers. To obtain the 3′ terminus of the viral genome, we used oligo$(dT)_{18}$ for cDNA synthesis, followed by amplification with consensus primers in the repeat (R) region of the long terminal repeat (LTR). This yielded a 742-bp fragment (fragment 12) that included the complete U3 and parts of the R region of the 3′ LTR. Sequences specific for the R region were then used to design primers to amplify the 5′ terminus (fragment 13). By this strategy, an entire genomic equivalent of SIVcpzTAN1 was obtained in 13 partially overlapping fragments. PCR products

TABLE 1. Oligonucleotide primers used to amplify SIVcpzTAN1 fragments comprising a complete genomic equivalent

Fragment[a]	Forward primer[b] (name; sequence)	Reverse primer[b,c] (name; sequence)	Size[d]
1	CON-1F1; CCAGCNCACAAAGGNATAGGAGG	CON-1R1; ACBACYGCNCCTTCHCCTTTC	650
	CON-1F1; CCAGCNCACAAAGGNATAGGAGG	CON-1R2; CCCAATCCCCCCTTTTCTTTTAAAATT	
2	CON-2F1; AAATGGCTGTGGTATATAAAAAT	CON-2R1; CCCWTCCAGTCCCCCCTTTTC	782
	CON-2F2; GCTTAAGAAAGGTTAGGCAGGG	CON-2R2; TCCCCCCTTTTCTTTTAAAAA	
3	CON-3F1; ATYTAYCARTAYATGGATGA	TAN1-3R1; ATTACTGTGATATTTATCATG	944
	CON-3F2; GGAAARYTAAATTGGGCAAGTC	TAN1-3R2; ATTTATCATGGTCTTCCTGTGC	
4	TAN1-4F1; CATACCAGAAGAAACAGGAA	CON-4R1; CTGGKGCTTGTTCCATCTATC	1,058
	TAN1-4F2; AATTGGCAGGAAGATGGCCTG	CON-4R1; CTGGKGCTTGTTCCATCTATC	
5	CON-5F1; CAGCGACARTAACGCTGACGG	TAN1-5R1; GCTGTTCTGGATCCTGCTCCGC	576
	CON-5F2; ACAATTATTGTCTGGTATAGTGCAACAGCA	TAN1-5R2; CGCTTGGGTAGGGATCTGTG	
6	CON-6F1; ATTAAAACCAGGAATGGATGG	TAN1-6R1; TATCTGATATGTCCACTGTCCT	948
	CON-6F2; CCAAGAGTAAARCAATGGCC	TAN1-6R2; GCCCCTGCTTCTGTATTGCTGC	
7	TAN1-7F1; AAGAATATTGACATTCTGCCACT	CON-7R1; TACTGCTTTGGTAYAGGATCTTGA	533
	TAN1-7F2; CAATCAGGACACAGACAGGTAGGGAC	CON-7R2; CGCTTCTTCCTGCCATAGGA	
8	CON-8F1; ACAGGAWGTRAAAAAYTGGATGAC	TAN1-8R1; GGTTTTCTGGTCCAACTGCAGA	849
	CON-8F2; ATGATGACAGCTTGTCAAGGAGT	TAN1-8R2; GTCTTACATATTTCTGTTAATGCC	
9	TAN1-9F1; GGAACAAGCACCGAACGACAAC	TAN1-9R1; GCACATCCCCAGAGGCTTAGG	2,197
	TAN1-9F2; CACTATCGCCATGGTTGTAT	TAN1-9R2; AGCTGTTGGTCTCTTATGT	
10	CON-10F1; CCMAGAACCTTAAATGCCTGGGT	TAN1-10R1; AGAATCTTTGCCTTATGAGAGGGT	494
	CON-10F2; RGAAGTMATCAATGAGGAAGC	TAN1-10R1; AGAATCTTTGCCTTATGAGAGGGT	
11	CON-11F1; AACCTGTGACGCCAGGACTCG	TAN1-11R1; CCTTGTGCCTGTGGGGTTGTC	752
	CON-11F2; GAAGATGGGTGCGAGAGCGTC	TAN1-11R2; CCTGCTTGCTGGGGTCCTG	
12	TAN1-12F1; AAGATCTTCATAACACCTCA	CON-12R1; AAGCTTTATTGAGCATTAAG	742
	TAN1-12F2; GGACAGGCCTCACAAGA	CON-12R2; TTGAGCATTAAGAAGCAGGGTCTCC	
13	TAN1-13F1; CTGGGAGCTCTCTGGTAGT	TAN1-13R1; ACGCTGTAGTTCGCTTCC	368
	TAN1-13F2; CTCTGGTAGTGGCTGGCT	TAN1-13R2; TTCCCATGTATCCAGCTTATCTCC	

[a] See Fig. 1 for position in SIVcpzTAN1 genome.
[b] CON primers were designed according to HIV-1-SIVcpz consensus sequences (20); TAN1 primers are strain specific. F1, first-round forward primer; F2, second-round forward primer; R1, first-round reverse primer; R2, second-round reverse primer. R = A/G, Y = C/T, M = A/C, K = G/T, S = C/G, W = A/T, H = A/C/T, B = C/G/T, V = A/C/G, D = A/G/T, and N = A/C/T/G.
[c] First-round reverse primers were also used for cDNA synthesis, except for fragment 12, where oligo(dT)$_{18}$ was used (see text for details).
[d] Size of nested (or seminested) PCR fragments (in base pairs).

were visualized by gel electrophoresis, purified (Qiagen, Valencia, Calif.), and sequenced without interim cloning, except for fragments 12 and 13, which were cloned into a plasmid (pTOPO) vector (Invitrogen, Carlsbad, Calif.) to obtain viral sequences immediately adjacent to the primer sequences. Viral sequences were assembled with Sequencher software (Gene Codes Corporation, Ann Arbor, Mich.). None of the directly sequenced fragments yielded ambiguous bases, suggesting amplification of viral sequences at limiting dilution conditions. However, a limited number of nucleotide sequence differences were noted between adjoining fragments in the region of sequence overlap, consistent with the presence of SIVcpz as a quasispecies in fecal samples (46). In these cases, the 5′ sequence was arbitrarily selected for compilation of the final SIVcpzTAN1 sequence. The concatenated SIVcpzTAN1 genome (R-U5-*gag*-*pol*-*env*-U3-R) was 9,326 bp in length.

Inspection of deduced protein sequences of SIVcpzTAN1 revealed the expected reading frames for *gag*, *pol*, *vif*, *vpr*, *vpu*, *tat*, *rev*, *env*, and *nef*, none of which contained inactivating mutations. In addition, major regulatory sequences, including promoter and enhancer elements in the LTR, the transactivation region stem-loop structure, the packaging signal, the primer binding site, and major splice sites all appeared to be intact. The deduced TAN1 Vpu amino acid sequence was highly divergent from those of other HIV-1 or SIVcpz Vpu proteins. For example, ANT and TAN1 Vpu proteins were only 37% identical (Fig. 2A). However, the position of the *vpu* open reading frame in the TAN1 genome (Fig. 1) and the hydrophobicity profile of its deduced protein sequence (Fig. 2B) were very similar to those of other HIV-1 and SIV strains (8, 41). In addition, secondary structure predictions suggested the presence of alpha helices near the C terminus that flanked two highly conserved serine residues (Fig. 2A) previously shown to be critical for HIV-1 Vpu-mediated CD4 degradation (41). Together, these data suggest that TAN1 encodes a functional Vpu protein.

To compare SIVcpzTAN1 to previously characterized SIVcpz and HIV-1 strains, we performed diversity plot analyses of concatenated protein sequences. Pairwise sequence distances were plotted for windows of 300 amino acids, moved in steps of 20 along the alignment. Figure 3 depicts distance plots illustrating the proportion of amino acid sequence differences between SIVcpzTAN1 and SIVcpzGAB1 (green solid line), U455 (HIV-1 group M; black solid line), MVP5180 (HIV-1 group O; blue solid line), and SIVcpzANT (red solid line), revealing an extent of sequence diversity between members of the HIV-1/SIVcpz group that is comparable to that observed for SIVs from pairs of closely related monkey species (the distances of SIVs from vervet and grivet monkeys, and from l'hoest and solatus monkeys, are shown for comparison as red and blue dotted lines, respectively; note that these viruses differ even less than the various SIVcpz and HIV-1 strains in the Env protein). The virus most similar to SIVcpzTAN1 was SIVcpzANT, but even this strain was quite divergent, with distance values ranging from 17% in Gag to 35% in Env. Phylogenetic analyses of Gag, Pol, and Env protein sequences

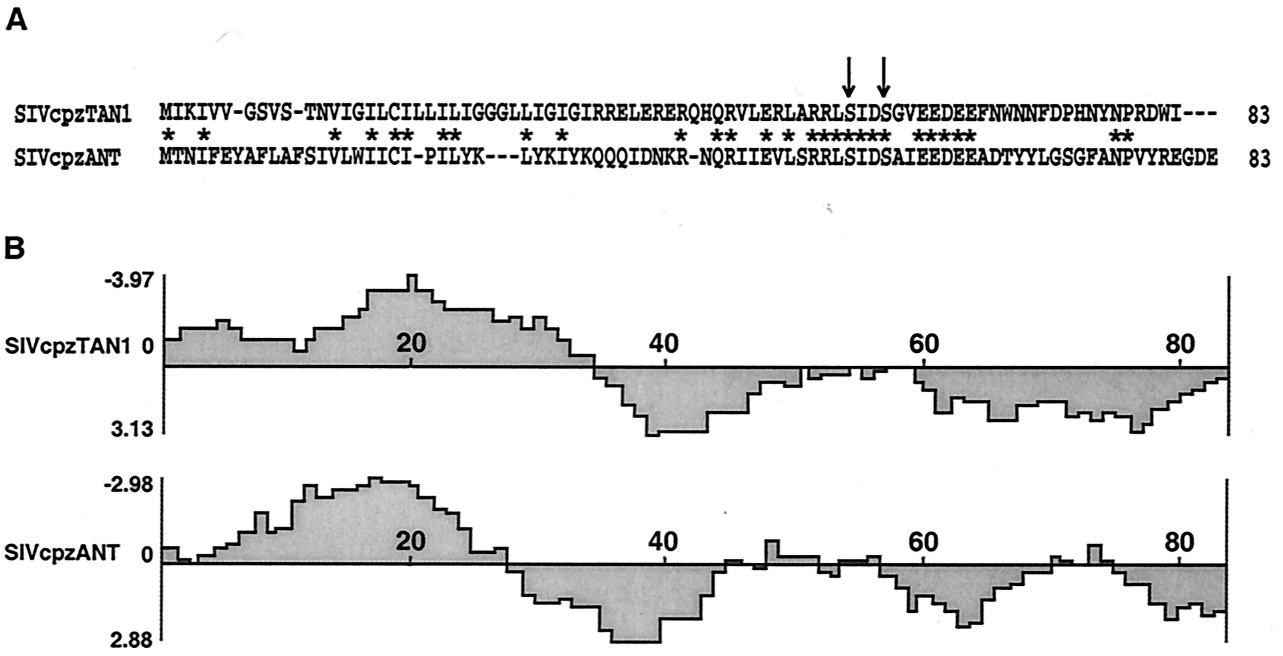

FIG. 2. Genetic diversity of Vpu in SIVcpz from *P. t. schweinfurthii*. (A) Alignment of the TAN1 and ANT Vpu sequences reveals an extraordinary degree of diversity (identical residues are indicated by asterisks) but demonstrates conservation of two serine residues (denoted by arrows) previously shown to be critical for HIV-1 Vpu-induced CD4 degradation (41). (B) Hydrophobicity profiles of SIVcpzTAN1 and SIVcpzANT Vpu proteins (calculated with the Kyte and Doolittle algorithm [21]) suggest similar secondary structures. Positive scores indicate degrees of hydrophobicity. The *x* axis shows the number of amino acid residues along the protein.

confirmed these relationships (Fig. 4), demonstrating that SIVcpzTAN1 fell within the HIV-1/SIVcpz radiation and grouped most closely with SIVcpzANT. This phylogenetic position was consistent in all major coding regions and supported by significant bootstrap values (Fig. 4). Distance and phylogenetic analyses thus identified SIVcpzTAN1 as a highly divergent member of the HIV-1/SIVcpz group of viruses. Since, until now, there has been only a single divergent *P. t. schweinfurthii* strain from a captive chimpanzee (Noah) of unknown origin, the possibility existed that SIVcpzANT was the result of a cross-species transmission event from another primate species and did not really represent a virus naturally infecting chimpanzees. The derivation of the complete SIVcpzTAN1 sequence from a chimpanzee of unquestionable provenance renders this possibility improbable. The phylogenetic position of TAN1 in Fig. 4 confirms the authenticity of SIVcpzANT as a bona fide SIVcpz strain and thus provides conclusive evidence for the existence of two major lineages within the SIVcpz/HIV-1 radiation.

We next inspected SIVcpz and HIV-1 protein alignments for lineage-specific amino acid sequence insertions and deletions because such features are excluded from phylogenetic analyses. This analysis identified several signatures that distinguished ANT and TAN1 from all other SIVcpz and HIV-1 strains (Fig. 5). For example, both ANT and TAN1 contained an identical 5-amino-acid insertion (KGPRR) near the C terminus of Vif which disrupted a highly conserved PPLP motif previously shown to be critical, in its entirety, for HIV-1 Vif function (38). In addition, they exhibited a 5-amino-acid deletion near the C terminus of Nef that included a diacidic β-COP (coatomer protein) binding motif shown elsewhere to be important for HIV-1 Nef-induced CD4 degradation (30). Both ANT and TAN1 also encoded a considerably truncated Vpr protein that lacked several basic residues at the C terminus previously shown to be important for HIV-1 Vpr-induced nuclear localization and G_2 cell cycle arrest, including a critical Arg-90 residue (9, 34, 47). Since accessory protein functions are highly conserved among divergent SIV lineages (19, 37, 40), it is highly unlikely that the Vif, Vpr, and Nef proteins of the two *P. t. schweinfurthii* viruses have lost these functions (this is especially true for TAN1, which was derived without the in vitro selection that might occur through growth in human T-cell lines). Instead, the observed Vif, Vpr, and Nef mutations are likely compensated by amino acid substitutions elsewhere in these proteins. Finally, both ANT and TAN1 exhibited an amino acid sequence insertion (8 amino acids for ANT; 9 amino acids for TAN1) in the ectodomain of the transmembrane envelope glycoprotein (gp41) which is flanked by two additional cysteine residues (Fig. 5). Unpaired cysteines are known to interfere with the proper folding of the SIV and HIV envelope glycoproteins (22, 42, 44). It is thus likely that the additional cysteine residues in TAN1 and ANT gp41 form intermolecular disulfide bonds, possibly resulting in an additional surface loop that might alter the local gp41 structure. Since this region is also known to be involved in gp120-gp41 interactions (5, 33), it is possible that compensatory changes in the N or C terminus of gp120 have evolved in association with these mutations. Construction of a replication-competent TAN1 genome and functional analyses of its gene products will allow us to examine these possibilities experimentally. Inter-

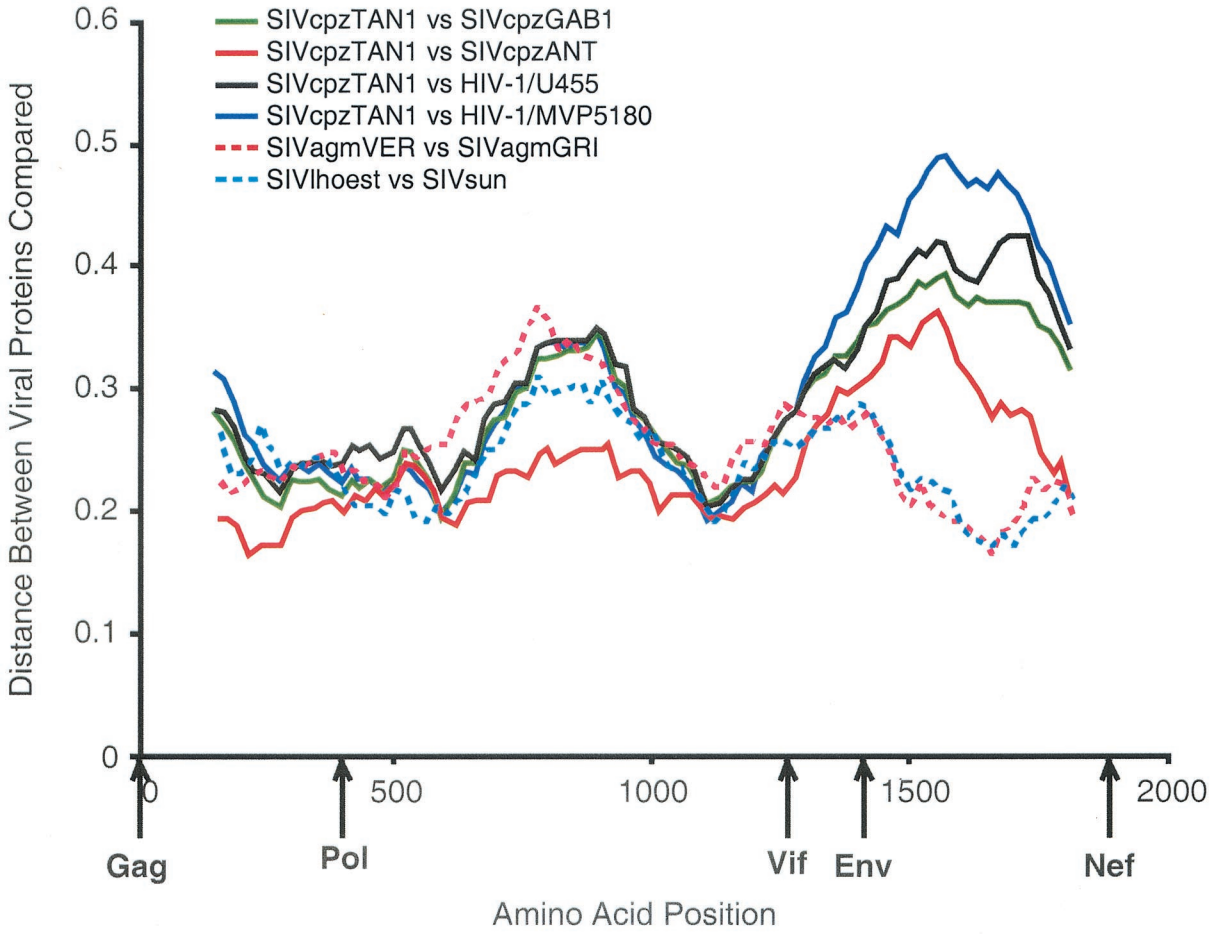

FIG. 3. Diversity plots of concatenated protein sequences illustrating the extent of genetic diversity within the HIV-1/SIVcpz group of viruses. The proportion of amino acid sequence differences between SIVcpzTAN1 and SIVcpzGAB1 (green solid line), SIVcpzANT (red solid line), HIV-1/U455 (black solid line), and HIV-1/MVP5180 (blue solid line) is shown, in comparison with the proportion of amino acid sequence differences between SIVagmVER155 and SIVagmGRI-1 (red dotted line) and SIVlhoest7 and SIVsun (blue dotted line), respectively. Predicted protein sequences were concatenated, with C-terminal overlaps being removed and aligned with Clustal W (43) and with minor manual adjustments being made with SEAVIEW (11). Sites that could not be unambiguously aligned, as well as sites containing a gap in any sequence, were excluded from the alignment. Protein sequence differences (number of amino acid sequence mismatches per sequence length) were calculated for windows of 300 amino acids, moved in steps of 20 amino acids across the concatenated alignment. The *x* axis indicates amino acid positions, with the N termini of Gag, Pol, Vif, Env, and Nef denoted. The *y* axis denotes the distance between the viral proteins (0.1 = 10% differences).

estingly, the extra cysteine pair in gp41, the truncated Vpr, and the Vif insertion were absent not only from SIVcpz from *P. t. troglodytes* but also from all other SIVs, including the relatively more closely related (at least in *env*) SIVgsn strain (8). This would suggest that *P. t. schweinfurthii* viruses acquired these changes some time after their divergence from the common SIVcpz ancestor but before the split of the lineages represented by today's SIVcpzTAN1 and SIVcpzANT. In addition, the absence of these signatures from all known HIV-1 variants (groups M, N, and O) is consistent with their west-central African chimpanzee (*P. t. troglodytes*) origin.

We previously interpreted the host-specific clustering of *P. t. troglodytes* and *P. t. schweinfurthii* viruses as indicating a long-standing virus-host relationship, possibly predating the divergence of these two chimpanzee subspecies hundreds of thousands of years ago (12, 15). Although the phylogenetic position of SIVcpzTAN1 is consistent with this interpretation, recent molecular clock calculations using more realistic models of HIV and SIV sequence evolution suggest a timescale for SIV evolution that is much shorter than previously assumed (36). For SIVcpz in particular, this estimated timescale would certainly not be on the order of hundreds of thousands or millions of years as required for virus-host coevolution. Moreover, mitochondrial DNA studies of wild chimpanzees in west-central Africa suggest that the two equatorial subspecies are not as distinct as previously believed (24), with *P. t. schweinfurthii* now being seen to fall within the *P. t. troglodytes* radiation (10). It is thus possible that SIVcpz was introduced into chimpanzees relatively more recently, most likely by cross-species transmission of SIV from smaller monkeys that serve as a food source for chimpanzees. The most probable location for such a transmission would be west-central Africa, since it is there where smaller monkey species have been determined to harbor related viruses (8, 13, 39). Once introduced, SIVcpz would have likely spread eastward to *P. t. schweinfurthii* but probably not to the western subspecies *Pan troglodytes verus* which is believed

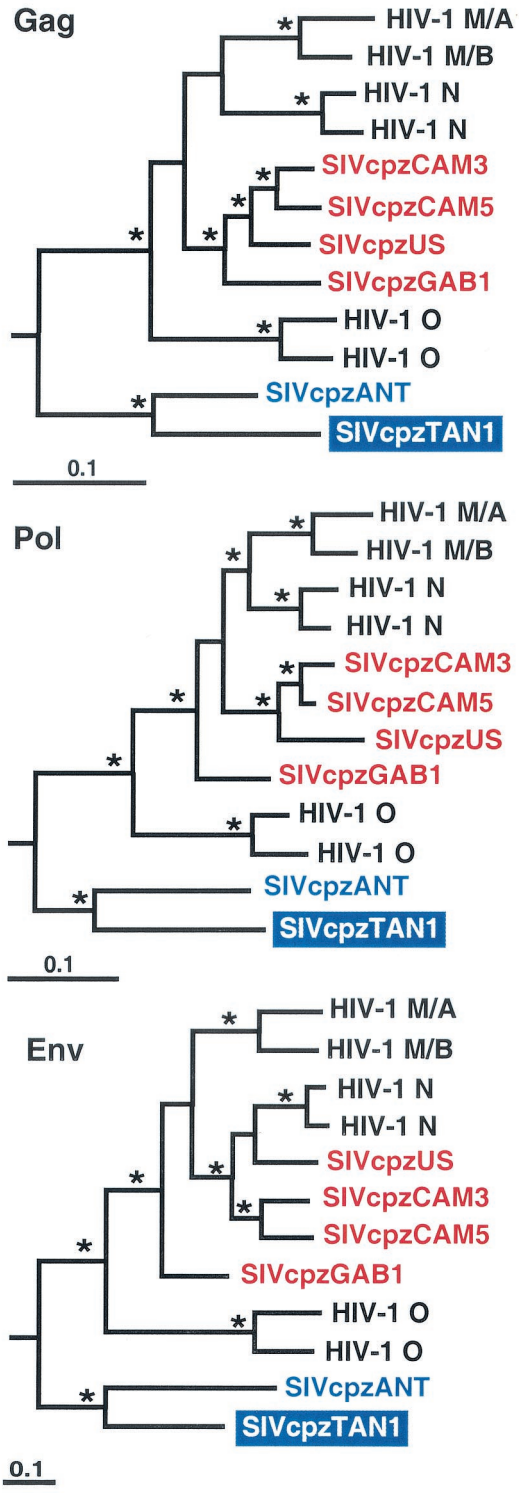

FIG. 4. Maximum likelihood tree depicting the relationship of SIVcpzTAN1 to other primate lentiviruses. SIVcpzTAN1 predicted protein sequences were compared to the following HIV and SIV representatives: HIV-1 group M subtype A (U455 [GenBank accession number M62320]) and subtype B (LAI [accession number K02013]), group N (YBF30 [accession number AJ006022] and YBF106 [accession number AJ271370]) and group O (ANT70 [accession number L20587] and MVP5180 [accession number L20571]), and SIVcpz (GAB1 [accession number X52154], US [accession number AF103818], CAM3 [accession number AF115393], CAM5 [accession number AJ271369], and ANT [accession number U42720]). Phylogenetic trees were constructed by both maximum likelihood and neighbor joining methods and rooted by using SIVmndGB1 as an outgroup (M27470). The neighbor joining method was implemented in Clustal W (43) with Kimura's empirical correction (18) and 1,000 bootstrapped replicates. The maximum likelihood method was implemented in PROTML from the MOLPHY package (J. Adachi and M. Hasegawa, Institute of Statistical Mathematics, Tokyo, Japan, 1994) by using a quick search and the Jones, Taylor, and Thornton model for amino acid substitution (16) with amino acid frequencies. SIVcpz strains from *P. t. troglodytes* are shown in red, and those from *P. t. schweinfurthii* are shown in blue, with SIVcpzTAN1 highlighted. Horizontal branch lengths are drawn to scale. The scale bar indicates 0.1 substitution per site. Asterisks denote branches (and thus the clades to their right) supported by more than 85% of bootstrapped replicates in trees made by the neighbor joining method.

to have been geographically isolated from *P. t. troglodytes* and *P. t. schweinfurthii* for an extended period of time (17, 24). The absence of SIVcpz from two wild *P. t. verus* communities in the Taï forest (32) as well as the lack of SIVcpz from over a thousand captive chimpanzees of west African origin (31; W. M. Switzer and W. Heneine, personal communication) strongly supports this conclusion. Thus, the two major SIVcpz lineages shown here to infect *P. t. troglodytes* and *P. t. schweinfurthii* subspecies (Fig. 4) most likely reflect a more recent divergence of SIVcpz within geographically isolated populations of chimpanzees rather than ancient infection of chimpanzees and coevolution of virus and host.

The analysis of the complete SIVcpzTAN1 sequence also provides new insight into the origins of HIV-1 groups M, N, and O. First, the significant clustering of SIVcpzTAN1 and SIVcpzANT, their divergence from SIVcpz strains infecting *P. t. troglodytes*, and the close phylogenetic relationship of *P. t. troglodytes* viruses to HIV-1 groups M, N, and O (Fig. 4) all reaffirm the conclusion that the HIV-1 epidemic resulted from zoonotic transmission of SIVcpz from west-central African and not east African chimpanzees to humans (12). Second, the discovery of SIVcpzTAN1 and its relationships to the other HIV-1 and SIVcpz strains provides a plausible explanation for the unusual position of HIV-1 group O in phylogenetic trees. Because group O viruses fall as an outgroup to (rather than within) the clade comprised of other HIV-1 groups and SIVcpz from *P. t. troglodytes*, a west-central African origin of HIV-1 group O has until now seemed less certain (12, 15, 35), even though epidemiologically group O viruses have long been known to be largely restricted to Cameroon and neighboring countries (1, 4, 27). Inspection of the phylogenetic trees in Fig. 4 and comparison of HIV-1 and SIVcpz protein sequence identities in Table 2 now reveal that, overall, HIV-1 group O viruses are no more distant from *P. t. troglodytes* viruses than the two *P. t. schweinfurthii* strains are from each other. Together with the known demographic distribution of HIV-1 group O infections, these data argue strongly for a west-central African chimpanzee origin of HIV-1 group O. It may then seem surprising that all known west-central African SIVcpz strains cluster together with HIV-1 groups M and N in a single clade, with no SIVcpz being closely related to HIV-1 group O. This may be due to the fact that all known *P. t. troglodytes* viruses, with the possible exception of SIVcpzUS, whose geographic origin is unknown, were derived from chimpanzees

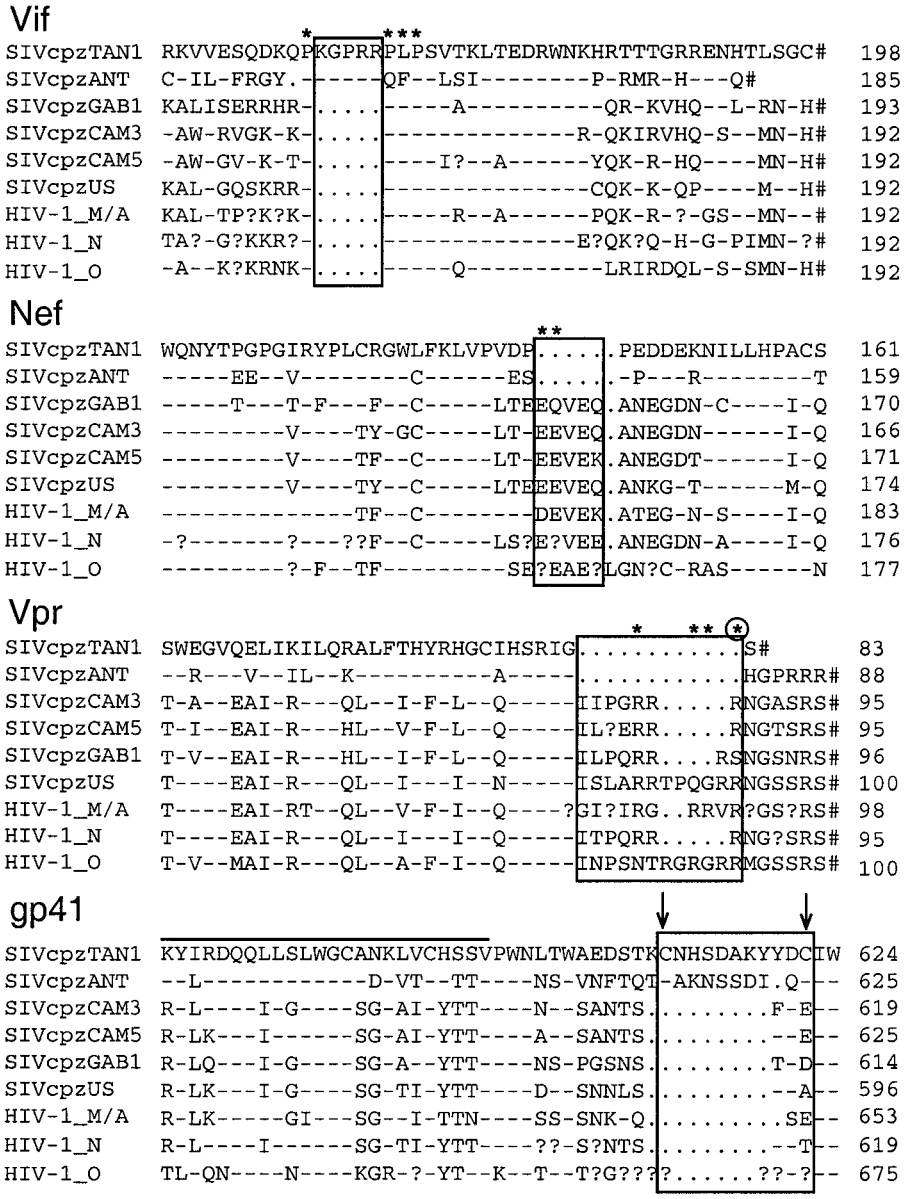

FIG. 5. SIVcpz lineage-specific protein signatures. Alignments of SIVcpz and HIV-1 Vif, Nef, Vpr, and gp41 deduced amino acid sequences are shown for selected regions of the proteins. Sequences are compared to SIVcpzTAN1, with dashes denoting sequence identity and dots representing gaps introduced to optimize the alignment. Question marks indicate sites of ambiguous sequence (in SIVcpz) or sites at which fewer than 50% of the viruses contain the same amino acid residue (in HIV-1). Pound signs indicate stop codons. HIV-1 group M, N, and O sequences are consensus sequences obtained from the Los Alamos HIV Sequence Database (http://hiv-web.lanl.gov) (20). For HIV-1 group M, consensus subtype A sequences are shown; however, consensus sequences from the other subtypes yielded the same results (data not shown). Vertical boxes highlight SIVcpz lineage-specific protein signatures in Vif, Vpr, Nef, and gp41. Arrows denote a pair of cysteine residues in the ectodomain of gp41 that is unique to *P. t. schweinfurthii* viruses (the horizontal line denotes the immunodominant region of the HIV-1 gp41 glycoprotein). Asterisks indicate a highly conserved PPLP motif in Vif, a diacidic β-COP binding motif in Nef, and four C-terminal Arg residues in Vpr (Arg-90 is circled), previously shown to be critical for protein function in HIV-1 (see the text for details). Consistent with their proposed origin (12), HIV-1 groups M, N, and O exhibit protein features characteristic of SIVcpz from *P. t. troglodytes*.

captured in southern Cameroon or northern Gabon in a region south of the Sanaga River and north of the Ogooué River (7, 28). Since HIV-1 groups M and N fall within the radiation of these chimpanzee viruses, it is likely that the cross-species transmission events giving rise to these HIV-1 groups took place in the same vicinity. The SIVcpz predecessor of HIV-1 group O would then be expected to have come from a different geographic region, yet still within west-central Africa. An intriguing possibility is that HIV-1 group O originated in the fourth (Nigerian) subspecies of chimpanzees, *Pan troglodytes vellerosus*, whose natural habitat includes eastern Nigeria and western Cameroon with the Sanaga River delineating its southern boundary (6, 10). Evidence of rare chimpanzee gene flow across the Sanaga River has been reported elsewhere (10),

TABLE 2. Protein sequence identities between different HIV-1 and SIVcpz groups of viruses

Comparison	Avg % (range) protein sequence identity								
	Gag	Pol	Vif	Vpr	Tat	Rev	Vpu	Env	Nef
SIVcpz (*P.t.t.*) vs SIVcpz (*P.t.t.*)[a]	83 (79–89)	90 (82–94)	74 (68–84)	83 (77–89)	68 (62–83)	64 (50–76)	46 (33–67)	69 (63–77)	80 (77–82)
HIV-1 M vs SIVcpz (*P.t.t.*)[b]	73 (71–75)	82 (80–85)	70 (66–74)	81 (74–88)	64 (56–76)	59 (48–67)	31 (20–40)	61 (59–64)	70 (65–76)
HIV-1 N vs SIVcpz (*P.t.t.*)[c]	75 (74–77)	83 (81–84)	72 (64–76)	83 (79–88)	69 (65–71)	62 (55–70)	37 (29–47)	69 (64–73)	72 (69–77)
HIV-1 O vs SIVcpz (*P.t.t.*)[d]	71 (70–72)	77 (76–79)	67 (62–69)	78 (75–82)	53 (49–57)	46 (41–50)	32 (28–35)	51 (50–52)	61 (58–62)
SIVcpz (*P.t.s.*) vs SIVcpz (*P.t.t.*)[e]	65 (63–68)	71 (70–73)	57 (53–61)	62 (59–65)	52 (49–57)	45 (38–53)	20 (16–26)	49 (46–52)	55 (53–57)
SIVcpzTAN1 vs SIVcpzANT	70	75	64	73	57	55	37	56	68

[a] SIVcpz viruses from *P. t. troglodytes* (*P.t.t.*) include GAB1, US, CAM3, and CAM5.
[b] HIV-1 group M includes U455, LAI, ELI, C2200, CF506, and V1850.
[c] HIV-1 group N includes YBF30 and YBF160.
[d] HIV-1 group O includes MVP5180 and ANT70.
[e] SIVcpz viruses from *P. t. schweinfurthii* (*P.t.s.*) include ANT and TAN1.

offering the possibility that SIVcpz could have spread from *P. t. troglodytes* to *P. t. vellerosus* in the past. Naturally infected *P. t. vellerosus* apes have not yet been identified, but only very few have been tested. Comprehensive screening of these chimpanzees and others throughout west-central Africa for SIVcpz infection is needed to test this hypothesis.

The documentation of SIVcpzTAN1 in Gombe extends the geographic range of SIVcpz from Gabon and Cameroon across equatorial Africa to Tanzania. Although additional field studies throughout west-central and east Africa will be required to understand fully the geographic distribution and prevalence of SIVcpz in *P. t. troglodytes* and *P. t. schweinfurthii*, the present results make it most likely that both of these subspecies are endemically infected. We considered the possibility that Ch-06 could have been infected by HIV-1 from an infected human, or by SIV from another primate species, but found both scenarios implausible. First, none of the Gombe chimpanzees has ever been held captive, relocated from a different part of Africa, or inoculated with human blood; even had they been exposed to blood or bodily secretions from an infected human, there are no known HIV-1 strains that are at all closely related to SIVcpzTAN1 (Fig. 4). Second, among over 30 African primate species known to harbor SIV, none have been found to carry viruses that are sufficiently closely related to SIVcpz to represent a viable direct source (15, 25). Third, to explain an infection of Ch-06 by an SIV from another primate species would require, on phylogenetic grounds, the same as yet unidentified primate reservoir to have transmitted similar viruses to both chimpanzees and humans on as many as 10 different occasions. Such a hypothetical primate species, whose habitat would also need to overlap the range of both *P. t. troglodytes* and *P. t. schweinfurthii*, is unlikely given the results of the rather extensive surveys of primates for HIV-1-related infections that have already been conducted (2, 25). Thus, human-to-chimpanzee transmission of an SIVcpz-like virus cannot be invoked to explain the existence of SIVcpzTAN1, and an additional concurrent primate reservoir is highly improbable.

There remain a number of questions regarding the prevalence, subspecies distribution, and evolutionary history of SIVcpz in chimpanzees. Although six additional SIVcpz-infected chimpanzees have recently been identified in Gombe National Park (M. L Santiago and B. H. Hahn, unpublished data), other communities in east Africa appear to be free of infection, indicating, overall, a surprisingly low prevalence of SIVcpz in east African chimpanzees. Moreover, even within an infected community like Kasekela, spread of SIVcpz from the single index case has not yet been documented (Santiago and Hahn, unpublished). This is in stark contrast to SIV infection of smaller primate species, where prevalence increases with age in wild troops and up to 90% of sexually active adults become infected (3, 25, 29). Additional field studies of wild-living chimpanzees are clearly needed to shed light on these and other issues. It is possible that studies of SIVcpz infection in wild chimpanzee communities will help explain not only the pathogenicity of HIV-1 in humans but also its current transmission rates, which have led to the AIDS pandemic. In a practical sense, the present study demonstrates that full-length SIVcpz sequences can be derived noninvasively and that blood sampling is not required to conduct comprehensive molecular epidemiological studies of primate lentiviruses in endangered species. Moreover, it has been shown that this noninvasive approach is applicable to other viral infections and primate populations (23). This should provide new opportunities for primatologists, conservationists, and biomedical scientists to combine their research efforts to study medically important pathogens and emerging infectious diseases.

Nucleotide sequence accession number. The GenBank accession number for the concatenated SIVcpzTAN1 genome sequence (R-U5-*gag-pol-env*-U3-R) is AF447763.

We thank John Moore, Michael Malim, Didier Trono, and Michael Emerman for helpful discussions; the staff at Gombe Stream Research Centre for collecting chimpanzee fecal samples; Julie Decker and Maria Salazar for technical assistance; and W. J. Abbott for artwork and manuscript preparation.

This work was supported in part by grants NO1 AI 85338, RO1 AI 44596, and RO1 AI 50529 (B.H.H. and P.M.S.); the UAB Center for AIDS Research Sequencing Core (P30 AI 27767); the Leakey Foundation and the National Science Foundation (M.N.M. and R.W.W.); the Jane Goodall Institute (S.K., A.E.P., and D.A.C.); and the Howard Hughes Medical Institute (G.M.S.).

REFERENCES

1. **Ayouba, A., P. Mauclere, P. M. Martin, P. Cunin, J. Mfoupouendoun, B. Njinku, S. Souquieres, and F. Simon.** 2001. HIV-1 group O infection in Cameroon, 1986 to 1998. Emerg. Infect. Dis. **7:**466–467.
2. **Bailes, E., R. R. Chaudhuri, M. L. Santiago, F. Bibollet-Ruche, B. H. Hahn, and P. M. Sharp.** 2002. The evolution of primate lentiviruses and the origins of AIDS, p. 65–96. *In* T. A. Leitner (ed.), The molecular epidemiology of human viruses. Kluwer Academic Publishers, Boston, Mass.
3. **Bibollet-Ruche, F., A. Galat-Luong, G. Cuny, P. Sarni-Manchado, G. Galat, J. P. Durand, X. Pourrut, and F. Veas.** 1996. Simian immunodeficiency virus infection in a patas monkey (*Erythrocebus patas*): evidence for cross-species transmission from African green monkeys (*Cercopithecus aethiops sabaeus*) in the wild. J. Gen. Virol. **77:**773–781.
4. **Bibollet-Ruche, F., M. Peeters, S. Mboup, E. Ekaza, R. Gandji, J. Torimiro, E. N. Mpoudi, J. Amblard, G. Dibanga, M. Saidou, E. Esu-Williams, M. Vanden Haesevelde, E. Saman, and E. Delaporte.** 1998. Molecular characterization of the envelope transmembrane glycoprotein of 13 new human immunodeficiency virus type 1 group O strains from six different African countries. AIDS Res. Hum. Retrovir. **14:**1281–1285.
5. **Binley, J. M., R. W. Sanders, B. Clas, N. Schuelke, A. Master, Y. Guo, F. Kajumo, D. J. Anselma, P. J. Maddon, W. C. Olson, and J. P. Moore.** 2000. A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure. J. Virol. **74:**627–643.
6. **Butynski, T. M.** 2001. Africa's great apes: an overview of current taxonomy, distribution, numbers, conservation status, and threats, p. 3–56. *In* B. B. Beck, T. S. Stoinski, M. Hutchins, T. L. Maple, B. Norton, A. Rowan, E. F. Stevens, and A. Arluke (ed.), Great apes and humans: the ethics of coexistence. Smithsonian Institution Press, Washington, D.C.
7. **Corbet, S., M. C. Muller-Trutwin, P. Versmisse, S. Delarue, A. Ayouba, J. Lewis, S. Brunak, P. Martin, F. Brun-Vezinet, F. Simon, F. Barre-Sinoussi, and P. Mauclere.** 2000. *env* sequences of simian immunodeficiency viruses from chimpanzees in Cameroon are strongly related to those of human immunodeficiency virus group N from the same geographic area. J. Virol. **74:**529–534.
8. **Courgnaud, V., M. Salemi, X. Pourrut, E. Mpoudi-Ngole, B. Abela, P. Auzel, F. Bibollet-Ruche, B. Hahn, A. M. Vandamme, E. Delaporte, and M. Peeters.** 2002. Characterization of a novel simian immunodeficiency virus with a *vpu* gene from greater spot-nosed monkeys (*Cercopithecus nictitans*) provides new insights into simian/human immunodeficiency virus phylogeny. J. Virol. **76:**8298–8309.
9. **Di Marzio, P., S. Choe, M. Ebright, R. Knoblauch, and N. R. Landau.** 1995. Mutational analysis of cell cycle arrest, nuclear localization and virion packaging of human immunodeficiency virus type 1 Vpr. J. Virol. **69:**7909–7916.
10. **Gagneux, P., M. K. Gonder, T. L. Goldberg, and P. A. Morin.** 2001. Gene flow in wild chimpanzee populations: what genetic data tell us about chimpanzee movement over space and time. Philos. Trans. R. Soc. Lond. B Biol. Sci. **356:**889–897.
11. **Galtier, N., M. Gouy, and C. Gautier.** 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. **12:**543–548.
12. **Gao, F., E. Bailes, D. L. Robertson, Y. Chen, C. M. Rodenburg, S. F. Michael, L. B. Cummins, L. O. Arthur, M. Peeters, G. M. Shaw, P. M. Sharp, and B. H. Hahn.** 1999. Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. Nature **397:**436–441.
13. **Georges-Courbot, M. C., C. Y. Lu, M. Makuwa, P. Telfer, R. Onanga, G. Dubreuil, Z. Chen, S. M. Smith, A. Georges, F. Gao, B. H. Hahn, and P. A. Marx.** 1998. Natural infection of a household pet red-capped mangabey (*Cercocebus torquatus torquatus*) with a new simian immunodeficiency virus. J. Virol. **72:**600–608.
14. **Goodall, J.** 1986. The chimpanzees of Gombe: patterns of behavior. Belknap Press, Cambridge, Mass.
15. **Hahn, B. H., G. M. Shaw, K. M. De Cock, and P. M. Sharp.** 2000. AIDS as a zoonosis: scientific and public health implications. Science **287:**607–614.
16. **Jones, D. T., W. R. Taylor, and D. W. Thornton.** 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. **8:**275–282.
17. **Kaessmann, H., V. Wiebe, and S. Paabo.** 1999. Extensive nuclear DNA sequence diversity among chimpanzees. Science **286:**1159–1162.
18. **Kimura, M.** 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge, United Kingdom.
19. **Kirchhoff, F., J. Munch, S. Carl, N. Stolte, K. Matz-Rensing, D. Fuchs, P. T. Haaft, J. L. Heeney, T. Swigut, J. Skowronski, and C. Stahl-Hennig.** 1999. The human immunodeficiency virus type 1 *nef* gene can to a large extent replace simian immunodeficiency virus *nef* in vivo. J. Virol. **73:**8371–8383.
20. **Kuiken, C., B. Foley, B. Hahn, P. Marx, F. McCutchan, J. Mellors, S. Wolinsky, and B. Korber.** 2001. HIV sequence compendium 2001. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, N.Mex.
21. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
22. **Lekutis, C., U. Olshevsky, C. Furman, M. Thali, and J. Sodroski.** 1992. Contribution of disulfide bonds in the carboxyl terminus of the human immunodeficiency virus type I gp120 glycoprotein to CD4 binding. J. Acquir. Immune Defic. Syndr. **5:**78–81.
23. **Ling, B., M. L. Santiago, S. Meleth, B. Gormus, H. M. McClure, C. Apetrei, B. H. Hahn, and P. M. Marx.** 2003. Noninvasive detection of new simian immunodeficiency virus lineages in captive sooty mangabeys: ability to amplify virion RNA from fecal samples correlates with viral load in plasma. J. Virol. **77:**2214–2226.
24. **Morin, P. A., J. J. Moore, R. Chakraborty, L. Jin, J. Goodall, and D. S. Woodruff.** 1994. Kin selection, social structure, gene flow, and the evolution of chimpanzees. Science **265:**1193–1201.
25. **Peeters, M., V. Courgnaud, B. Abela, P. Auzel, X. Pourrut, F. Bibollet-Ruche, S. Loul, F. Liegeois, C. Butel, D. Koulagna, E. Mpoudi-Ngole, G. M. Shaw, B. H. Hahn, and E. Delaporte.** 2002. Risk to human health from a plethora of simian immunodeficiency viruses in primate bushmeat. Emerg. Infect. Dis. **8:**451–457.
26. **Peeters, M., K. Fransen, E. Delaporte, M. Van den Haesevelde, G. M. Gershy-Damet, L. Kestens, G. van der Groen, and P. Piot.** 1992. Isolation and characterization of a new chimpanzee lentivirus (simian immunodeficiency virus isolate cpz-ant) from a wild-captured chimpanzee. AIDS **6:**447–451.
27. **Peeters, M., A. Gueye, S. Mboup, F. Bibollet-Ruche, E. Ekaza, C. Mulanga, R. Ouedrago, R. Gandji, P. Mpele, G. Dibanga, B. Koumare, M. Saidou, E. Esu-Williams, J. P. Lombart, W. Badombena, N. Luo, M. Vanden Haesevelde, and E. Delaporte.** 1997. Geographical distribution of HIV-1 group O viruses in Africa. AIDS **11:**493–498.
28. **Peeters, M., C. Honore, T. Huet, L. Bedjabaga, S. Ossari, P. Bussi, R. W. Cooper, and E. Delaporte.** 1989. Isolation and partial characterization of an HIV-related virus occurring naturally in chimpanzees in Gabon. AIDS **3:**625–630.
29. **Phillips-Conroy, J. E., C. J. Jolly, B. Petros, J. S. Allan, and R. C. Desrosiers.** 1994. Sexual transmission of SIVagm in wild grivet monkeys. J. Med. Primatol. **23:**1–7.
30. **Piguet, V., F. Gu, M. Foti, N. Demaurex, J. Gruenberg, J. L. Carpentier, and D. Trono.** 1999. Nef-induced CD4 degradation: a diacidic-based motif in Nef functions as a lysosomal targeting signal through the binding of beta-COP in endosomes. Cell **97:**63–73.
31. **Prince, A. M., B. Brotman, D. H. Lee, L. Andrus, J. Valinsky, and P. Marx.** 2002. Lack of evidence for HIV type 1-related SIVcpz infection in captive and wild chimpanzees (*Pan troglodytes verus*) in west Africa. AIDS Res. Hum. Retrovir. **18:**657–660.
32. **Santiago, M. L., C. M. Rodenburg, S. Kamenya, F. Bibollet-Ruche, F. Gao, E. Bailes, S. Meleth, S. J. Soong, J. M. Kilby, Z. Moldoveanu, B. Fahey, M. N. Muller, A. Ayouba, E. Nerrienet, H. M. McClure, J. L. Heeney, A. E. Pusey, D. A. Collins, C. Boesch, R. W. Wrangham, J. Goodall, P. M. Sharp, G. M. Shaw, and B. H. Hahn.** 2001. SIVcpz in wild chimpanzees. Science **295:**465.
33. **Schulz, T. F., B. A. Jameson, L. Lopalco, A. G. Siccardi, R. A. Weiss, and J. P. Moore.** 1992. Conserved structural features in the interaction between retroviral surface and transmembrane glycoproteins? AIDS Res. Hum. Retrovir. **8:**1571–1580.
34. **Selig, L., S. Benichou, M. E. Rogel, L. I. Wu, M. A. Vodicka, J. Sire, R. Benarous, and M. Emerman.** 1997. Uracil DNA glycosylase specifically interacts with Vpr of both human immunodeficiency virus type 1 and simian immunodeficiency virus of sooty mangabeys, but binding does not correlate with cell cycle arrest. J. Virol. **71:**4842–4846.
35. **Sharp, P. M., E. Bailes, R. R. Chaudhuri, C. M. Rodenburg, M. O. Santiago, and B. H. Hahn.** 2001. The origins of acquired immune deficiency syndrome viruses: where and when? Philos. Trans. R. Soc. Lond. B Biol. Sci. **356:**867–876.
36. **Sharp, P. M., E. Bailes, F. Gao, B. E. Beer, V. M. Hirsch, and B. H. Hahn.** 2000. Origins and evolution of AIDS viruses: estimating the time-scale. Biochem. Soc. Trans. **28:**275–282.
37. **Simon, J. H., D. L. Miller, R. A. Fouchier, M. A. Soares, K. W. Peden, and M. H. Malim.** 1998. The regulation of primate immunodeficiency virus infectivity by Vif is cell species restricted: a role for Vif in determining virus host range and cross-species transmission. EMBO J. **17:**1259–1267.

38. **Simon, J. H., A. M. Sheehy, E. A. Carpenter, R. A. Fouchier, and M. H. Malim.** 1999. Mutational analysis of the human immunodeficiency virus type 1 Vif protein. J. Virol. **73:**2675–2681.
39. **Souquiere, S., F. Bibollet-Ruche, D. L. Robertson, M. Makuwa, C. Apetrei, R. Onanga, C. Kornfeld, J. C. Plantier, F. Gao, K. Abernethy, L. J. White, W. Karesh, P. Telfer, E. J. Wickings, P. Mauclere, P. A. Marx, F. Barre-Sinoussi, B. H. Hahn, M. C. Muller-Trutwin, and F. Simon.** 2001. Wild *Mandrillus sphinx* are carriers of two types of lentivirus. J. Virol. **75:**7086–7096.
40. **Stivahtis, G. L., M. A. Soares, M. A. Vodicka, B. H. Hahn, and M. Emerman.** 1997. Conservation and host specificity of Vpr-mediated cell cycle arrest suggest a fundamental role in primate lentivirus evolution and biology. J. Virol. **71:**4331–4338.
41. **Strebel, K.** 1996. Structure and function of HIV-1 Vpu, p. III19–III27. *In* G. Myers, B. Korber, B. Foley, K. T. Jeang, J. W. Mellors, and S. Wain-Hobson (ed.), Human retroviruses and AIDS 1996. Los Alamos National Laboratory, Los Alamos, N.Mex.
42. **Syu, W. J., W. R. Lee, B. Du, Q. C. Yu, M. Essex, and T. H. Lee.** 1991. Role of conserved gp41 cysteine residues in the processing of human immunodeficiency virus envelope precursor and viral infectivity. J. Virol. **65:**6349–6352.
43. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
44. **Tschachler, E., H. Buchow, R. C. Gallo, and M. S. Reitz, Jr.** 1990. Functional contribution of cysteine residues to the human immunodeficiency virus type 1 envelope. J. Virol. **64:**2250–2259.
45. **Vanden Haesevelde, M. M., M. Peeters, G. Jannes, W. Janssens, G. van der Groen, P. M. Sharp, and E. Saman.** 1996. Sequence analysis of a highly divergent HIV-1-related lentivirus isolated from a wild captured chimpanzee. Virology **221:**346–350.
46. **van der Hoek, L., R. Boom, J. Goudsmit, F. Snijders, and C. J. Sol.** 1995. Isolation of human immunodeficiency virus type 1 (HIV-1) RNA from feces by a simple method and difference between HIV-1 subpopulations in feces and serum. J. Clin. Microbiol. **33:**581–588.
47. **Zhou, Y., Y. Lu, and L. Ratner.** 1998. Arginine residues in the C-terminus of HIV-1 Vpr are important for nuclear localization and cell cycle arrest. Virology **242:**414–424.